



American Society of  
Agricultural and Biological Engineers

**An ASABE Meeting Presentation**

**Paper Number: 072101**

## **Sensitivity Analysis And Interdependence Of The SWAT Model Parameters**

**Remegio B. Confesor Jr., Research Associate** (confesor@onid.orst.edu)

Oregon State University, 3450 SW Campus Way, Corvallis, OR 97331

**Gerald Whittaker, Research Hydrologist**

NFSPRC-USDA-ARS, 3450 SW Campus Way, Corvallis, OR 97331

**Written for presentation at the  
2007 ASABE Annual International Meeting  
Sponsored by ASABE  
Minneapolis Convention Center  
Minneapolis, Minnesota  
17 - 20 June 2007**

**Abstract.** *In contrast to lumped-parameter models, the distributed and processed-based hydrologic models take into account the spatial distribution of the hydrologic processes but became highly parameterized. In the Soil and Water Assessment Tool (SWAT) for example, the watershed is subdivided into spatial units (subbasins and hydrologic response units, HRU's) and each spatial unit has its own unique parameters that are utilized in SWAT simulation. Sensitivity analyses had been used as screening tools for reducing the number of parameters in model calibration. The objective of this study was to analyze the sensitivity of the objective functions to changes in parameters used in the multiobjective automatic calibration of the SWAT model. We used a Bayesian network to estimate the interdependencies of the SWAT parameters. The direct and indirect effect of the parameters on the model output was also explored. Where there are multiple objectives, the parameters and their interaction in searching for the Pareto optimum change with position along the Pareto front. The information derived from the Bayesian network requires redefining sensitivity to include a description of the interaction of parameters in the calibration search process.*

**Keywords.** SWAT, Bayesian network, sensitivity, automatic calibration, Pareto, genetic algorithm.

## Introduction

Over several decades, hydrologic modeling has evolved from simple methods including empirical formulas, unit hydrographs, and analytical equations (i.e., Loague and Freeze, 1985) into conceptual rainfall-runoff such as the Sacramento Soil Moisture Accounting (SAC-SMA) and distributed process-based models such as the Soil and Water Assessment Tool (SWAT) and Hydrologic Simulation Program-Fortran (HSPF). Consequently, the simulation output has expanded from single runoff events to continuous time series that eventually included streamflow and water quality parameters. In contrast to lumped-parameter models, the distributed and processed-based models take into account the spatial distribution of the hydrologic processes but became highly parameterized. In SWAT for example, the watershed is subdivided into spatial units (subbasins and hydrologic response units, HRU's) and each spatial unit has its own unique parameters that are utilized in the simulation.

Sensitivity analyses have tended to be used as screening tools for reducing the number of parameters used in calibration. In an one at time (OAT) approach to sensitivity, each parameter is perturbed and the objective of the calibration is recalculated. Information about the effect of the parameter on calibration and the existence of interactions with other parameters is obtained, but no specific interactions are examined. Without information on the interaction of the parameters that produced the simulation, it is difficult to convey if the simulation, though accurate, makes physical sense. Previous studies acknowledged the correlations and interdependence between model parameters (Vrugt et al. 2006; Feyen et al., 2006). The Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Freer, 2001) methodology analyzes sets of parameters, and purports to take account of interactions among parameters, but no specific information about the interactions is produced by the method.

Sensitivity analysis can be further applied in the building, use, and understanding of models (Ho et al., 2005). Tarantola and Saltelli (2003) suggested that sensitivity analysis could obtain vital information about the simulated system; such as: identification of calibration variables to model reduction or simplification, better understanding of the model structure for given components of a system, model quality assurance, and model building in general. To date, no one has yet developed a method or reported the relationships and interdependence of parameters in hydrologic model calibration and simulation.

Bayesian networks learn from data the model's structure and the local distributions' parameters and have been used in gene regulatory networks, medical research, text analysis, and image processing. In this study, we explored the application of Bayesian networks in the sensitivity analysis of the SWAT model. Specifically, we analyzed the sensitivity of the objective functions to changes in parameters used in the multiobjective automatic calibration of the SWAT model.

## Bayesian Networks

A Bayesian network is a graphical model for probabilistic relationships among a set of variables. When used in conjunction with statistical techniques, this type of graphical model can be used to learn causal relationships. Bayesian networks have been applied in many fields (see Lauritzen et al., (2003) for an overview), but have not been utilized in hydrological modeling. An integrated model of fish population and their environment that was developed using a Bayesian network (Borsuk et al., 2006) is the closest application reported as of this date.

The Bayesian network applied in this study is defined by a directed acyclic graph (DAG), where directed means that there is a direction for the links among variables, and acyclic precludes a set of variables linked in a loop. The Bayesian network is learned by searching for all possible

combinations of links among variables and each combination is scored. Even when constrained to directed and acyclic graphs, it is easy to see that the number of possible combinations presents a formidable computational challenge as the number of variables increase. Where undirected cycles are allowed, exact inferences quickly become intractable.

We use a method of learning a Bayesian network developed by Bøttcher (2001). The method relies on the assumption of conditional Gaussian networks. It is further assumed that the parameters associated with one variable (e.g., mean, variance, and coefficients) are independent of the parameters associated with other variables, denoted as global parameter independence. Although Bayesian networks lend themselves to inference with incomplete data, here we assume that the data is complete, an assumption that is also relied on in the calibration procedure.

In order to determine which DAG is selected to represent the conditional dependencies among a set of random variables, expert knowledge can be used, or a network score calculated based on how well a DAG represents the conditional dependencies. The posterior probability of the DAG,  $p(d | D)$  is sometimes used to score a DAG ( $D$ ) given data ( $d$ ), from Bayes theorem:

$$p(D | d) \propto p(d | D)p(D),$$

where,  $p(d | D)$  is the likelihood of  $D$  and  $p(D)$  is the prior probability. The network score used here is the relative probability,

$$p(D, d) = p(d | D)p(D).$$

The network is learned by finding the DAG with the highest network score. As the number of variables increases, the calculation of scores for all possible networks becomes infeasible. Bøttcher's method uses a greedy search among networks with random restarts as a strategy that makes it unnecessary to score all possible networks. The greedy search compares two networks that differ by one arrow (connection) and selects the one with highest score then proceeds to the next comparison. To avoid local minima, the starting structure is perturbed and the search is restarted. The final network is selected from the networks resulting from the restarted searches.

## Methods

### *Model Description*

The Soil and Water Assessment Tool (SWAT; Arnold et al., 1998) was developed by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) "to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over a long period of time." SWAT is physically based, uses readily available inputs, is computationally efficient, and is a continuous model that operates on a daily time step. SWAT is not designed to simulate single-event storms. The buildup of pollutants and their impact on water bodies can be studied with SWAT simulation runs spanning over several decades. In SWAT, the entire watershed can be divided into several subbasins and each subbasin is further divided into unique combinations of land use and soil properties called the Hydrologic Response Unit (HRU). However, the location of each HRU is not specified in the subbasin. The Geographic Information System (GIS) interface (AVSWAT2000) is usually used to input and designate land use, soil, weather, groundwater, water use, management, pond and stream water quality data,

and the simulation period (Di Luzio et al., 2001). GIS input files include digital elevation model (DEM), land use and soil properties layers, and weather database.

### ***Watershed Description***

The Calapooia river watershed (US Geological Survey, USGS, 10 digit HUC 1709000303) is a tributary of the Willamette river basin west of the Cascades mountain range in Oregon (Figure 1). It has drainage area of 963 km<sup>2</sup> as delineated from a USGS streamflow gaging station (44°37'15" N, 123°07'40" W) in Albany, Linn County, Oregon. Its elevation ranged from 56 m to 1576 m and its land use is mainly agriculture (43%), forest (41.8%), and hay/pasture/range areas (11.2%). The remaining areas were composed of wetlands, urban areas, and water bodies. The 10-m DEM used in delineating the watersheds was taken from the Regional Ecosystem Office ([http://www.reo.gov/reo/data/DEM\\_Files/indexes/orequadindex.asp](http://www.reo.gov/reo/data/DEM_Files/indexes/orequadindex.asp)). The observed daily streamflow data used in calibrating SWAT were obtained from the USGS National Water Information System (NWIS) website (<http://nwis.waterdata.usgs.gov/nwis/discharge>). The state soil geographic (STATSGO) database for Oregon was from the US Department of Agriculture - National Resources Conservation Service, USDA-NCRS (<http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo>). Land use for the Willamette basin was acquired from the USGS National Water-Quality Assessment (NAWQA) Program ([http://or.water.usgs.gov/projs\\_dir/pn366/landuse.html](http://or.water.usgs.gov/projs_dir/pn366/landuse.html)). Climate data were taken from the Oregon Climatic Service (<http://www.ocs.oregonstate.edu/>).

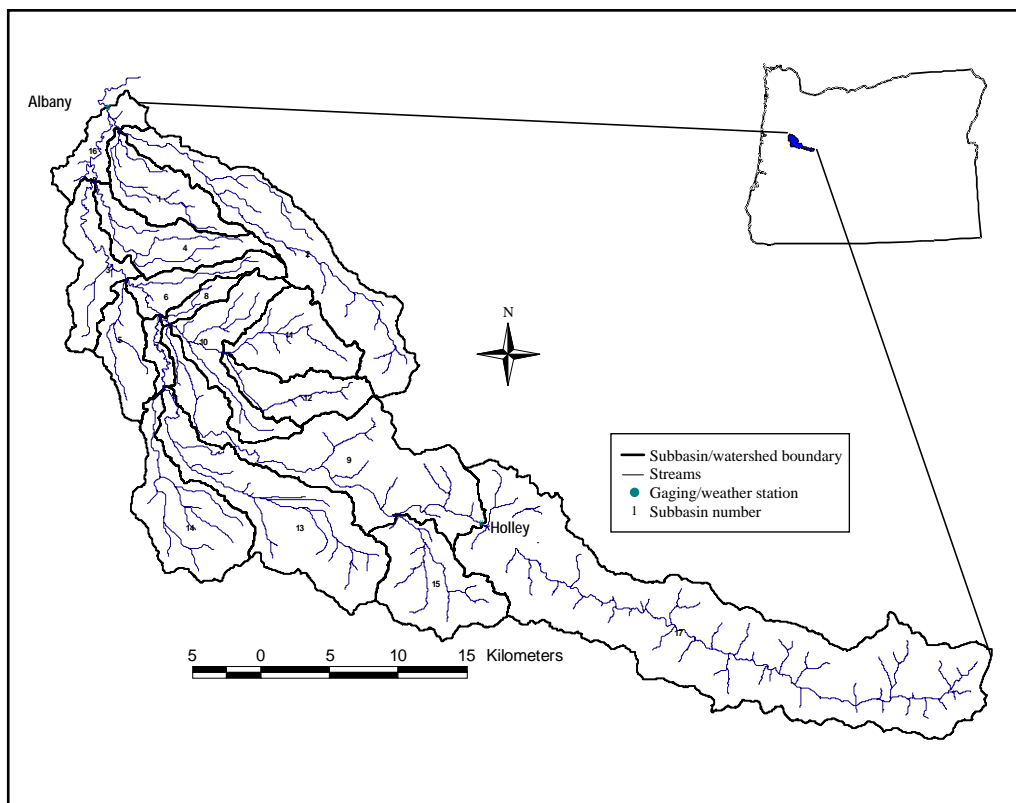


Figure 1. The Calapooia river watershed (delineated into 17 subbasins) in Oregon, USA.

## Automatic Calibration

The autocalibration method proposed by Confesor and Whittaker (2007) was employed in this study. The SWAT model was initially set up using the Arcview interface (AVSWAT2000) to SWAT (Di Luzio et al., 2001). The subbasins were delineated with a threshold size of 2100 hectares and the dominant landuse method, resulting in 17 subbasins. Based on the SWAT user's manual (Neitsch et al., 2002) and previous SWAT sensitivity analysis studies (Eckhardt and Arnold, 2001; Van Liew et al., 2005), sixteen variables were used in the calibration. The 17 subbasins were grouped into 3 locations representing steep, medium, and flat areas of the watershed. The subbasin/HRU slope (HRUSLP) and slope length (SLSUB) were optimized for each location group in the calibration. The curve number (CN) for each unique landuse and hydrologic group combination was explicitly calibrated resulting in 5 CN parameters. The soil evaporation compensation factor (ESCO) was also calibrated for each dominant landuse. This scheme produced 25 parameters to be optimized in the autocalibration. The limits of the variables for calibration were fixed to ensure realistic and acceptable values representative of the watershed characteristics. The calibration (October 1, 1972 to September 30, 1976) and validation (October 1, 1976 to September 30, 1980) periods were set for four water years.

The objective functions were to minimize the average Root Mean Square Error (RMSE) of the observed vs. simulated peak (driven) flows and to minimize the average RMSE of the observed vs. simulated low (non-driven) flows. The RMSE was defined as:

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2 \right]^{0.5}$$

where,  $n$  is the number of time steps with peak or low flow events,  $Q_{obs,i}$  is the observed streamflow at time  $i$ , and  $Q_{sim,i}$  is the simulated streamflow at time  $i$ .

The hydrographs were partitioned into driven and nondriven components assuming that the behavior of the watershed is different during the periods driven by rainfall and periods without rain (Boyle et al., 2001). The driven flow can be associated with the rising limb of the hydrograph and the nondriven flow with the recession flow. A baseflow filter was used to estimate the baseflow component of the observed streamflow (Arnold et al., 1995; Arnold and Allen, 1999). The streamflow was designated as driven when the first pass baseflow was less than 80% of the observed streamflow; otherwise the streamflow was classified as nondriven. The Nash-Sutcliffe model efficiency (NSE) was used to evaluate SWAT's overall performance at calibration and validation:

$$NSE = \left[ 1 - \frac{\frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2} \right]$$

where,  $\bar{Q}_{obs}$  is average of the observed daily flows and all the other variables are as previously defined. The Nash-Sutcliffe efficiency ranges from negative infinity to 1, with 1 indicating a perfect fit.

The implementation of the multiobjective evolutionary algorithm was simplified by using the genetic algorithm package (genalg) of the R statistical language (<http://www.r-project.org>). The computational scheme with two objective functions used in this study was implemented in a Beowulf cluster with 24 nodes (see Confesor and Whittaker, 2007) and is shown in Figure 2.

SWAT and the nondominated sorting genetic algorithm's (NSGA; Deb et al., 2002) nondomination ranking were created in Fortran as shared libraries callable in R. An initial population of 1000 solutions was randomly generated. Each solution contains the values of the 25 calibrated parameters assigned in real-coded string. SWAT was called as a subroutine and its source code was modified so that the values of the parameters of each solution were read instead of the values from the input files previously generated by AVSWAT2000. The SWAT daily streamflow output was used to evaluate each solution with the objective functions.

In the first iteration, a child population (size=1000) was then generated through selection, crossover, and mutation of the initial parent population. SWAT was then called and the child population was also evaluated with the objectives functions. NSGA ranking was called as a subroutine to implement non-dominated sorting on the combined parent and child population. The best 1000 individuals were selected as the next parent population to generate the new child population for the next iteration using crossover and mutation. After the first iteration, only the new child population is evaluated for the two objective functions since the parent population was previously evaluated and its fitness was already known. The optimization was stopped at 500 generations because changes in the objective functions' values were very small.

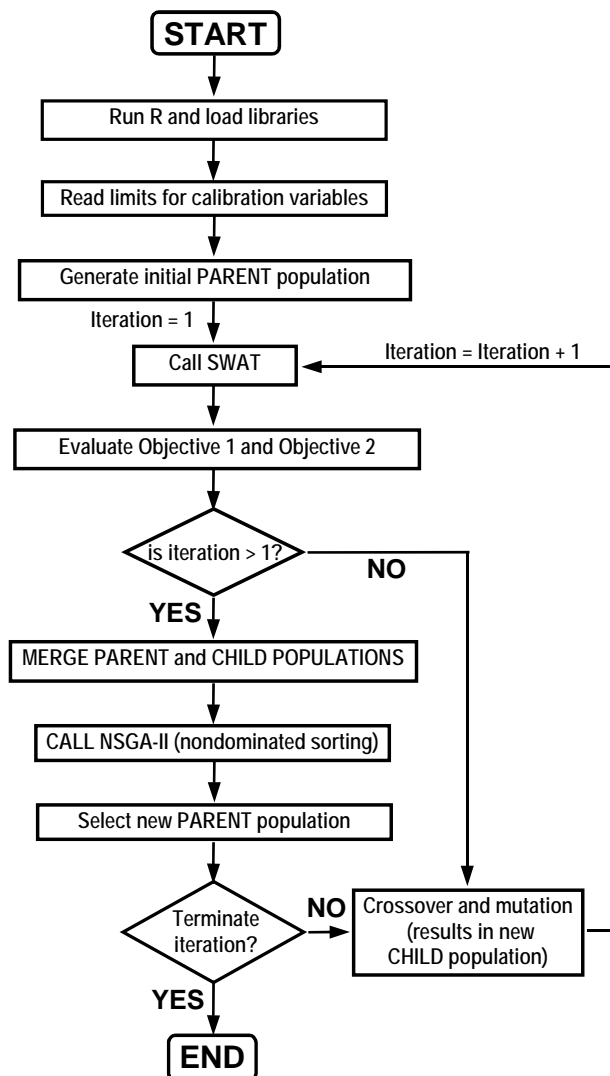


Figure 2. Genetic Algorithm Computational Scheme in R Linking NSGA-II and SWAT with 2 Objective Functions.

## Implementation of Bayesian networks

We hypothesized that different networks were used in different parts of the Pareto front, based on the observation that different parameters affect event driven flow and nondriven flow (the two objectives) in different ways. To examine this hypothesis, we extracted three subsets of 50 calibrated models each from the Pareto front after 500 generations (Figure 3). These subsets represent the solutions with: 1) the lowest event driven RMSE, 2) the objectives given approximately the same weight, and 3) the lowest nondriven flow RMSE. The Bayesian networks were then constructed from each of the three groups. To check that the network represents the calibrated model, one thousand (1000) solutions were simulated from each of the Bayesian network. The simulation is independent of the data used in the calibration, and consists of draws from the Gaussian distribution associated with each parameter and the regression relationships among parameters. These 1000 solutions were then evaluated by running the SWAT model and the corresponding objective functions calculated. We used the Deal R package (Bøttcher and Dethlefsen, 2003) for implementing estimation of the Bayesian network described above. All calculations were programmed and run in R. The Rgraphviz package was used for visualization of the networks (<http://www.bioconductor.org/repository/release1.5/package/html/Rgraphviz.html>).

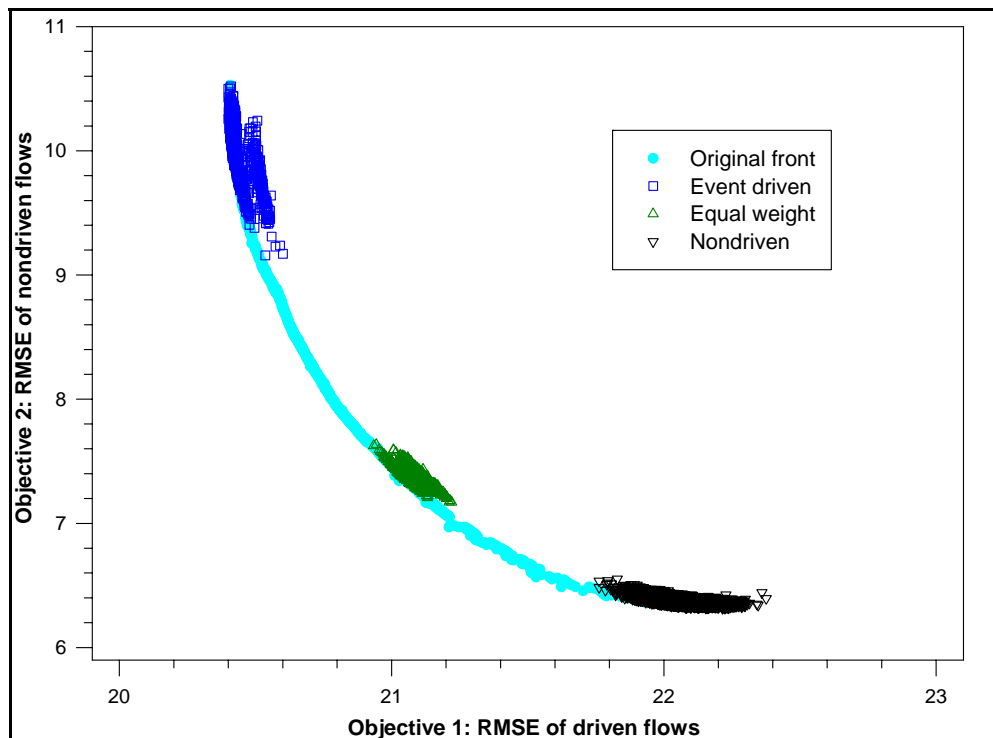


Figure 3. The objective space showing the Pareto front (500th iteration) and the solutions generated with the Bayesian network at different parts of the front.

## One At a Time (OAT) sensitivity

The OAT method was applied to the calibrated models as a comparison to the Bayesian network results. The method is a simple variation of calculation of a numerical partial derivative, where a single parameter is perturbed and the effect on the objective function observed. Given

a  $p$  dimensional vector of parameters ( $\alpha$ ), the sensitivity of the  $i$ th parameter to a perturbation  $\Delta$  is

$$d(\alpha_i | \alpha) = \frac{[y(\alpha_i, \dots, \alpha_{i-1}, \alpha_i + \Delta, \alpha_{i+1}, \dots, \alpha_p) - y(\alpha)]}{\Delta},$$

where,  $y(\alpha)$  is the model output of interest. In this method, the effect is conditioned on the level of the other parameters, and no interaction with other parameters is considered. Each of the 25 parameters was perturbed in all the 1000 solutions of the 500th generation. The sensitivity of each parameter was then calculated as the average of 1000 solutions.

## Discussion

### Calibration and Validation

The Pareto front consisting of 1000 solutions at the 500th iteration is shown in Figure 3. Each solution in the front was a calibrated model with a unique combination of the 25 parameters. The daily Nash-Sutcliffe model efficiency for these calibrated models varied between 0.81 and 0.87. There was no correlation between position on the front and the model efficiency measure. The daily Nash-Sutcliffe model efficiency for the validated models ranged from 0.76 to 0.81. Figure 4 shows the range of the simulated flows with the observed value at both the calibration and validation periods.

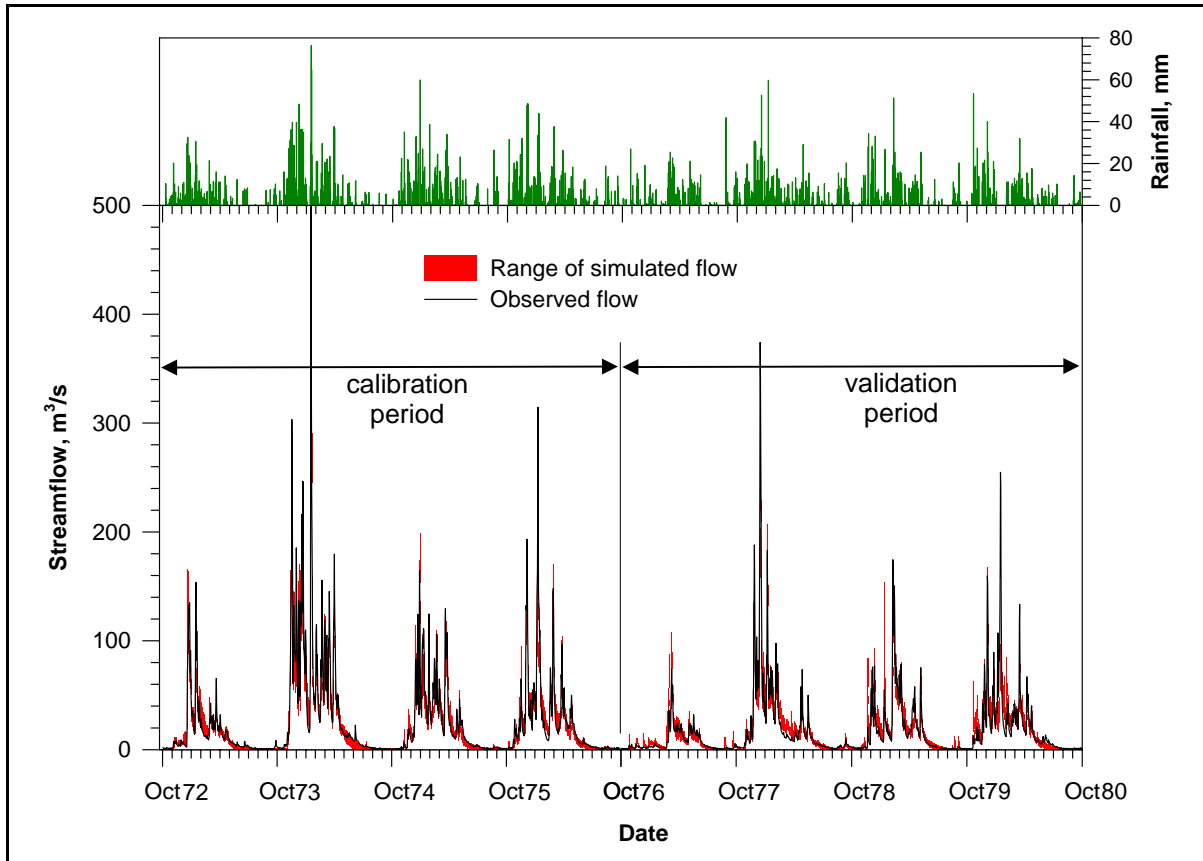


Figure 4. Observed streamflow and the simulated flow at both the calibration and validation periods, in the Calapooia watershed, OR, USA.

## Sensitivity

In its simplest sense, sensitivity is the change in a single objective function in response to change in a parameter influencing the objective function. Table 1 presents the result of a simple one at a time perturbation calculation of sensitivity for the two objectives in this study. It should be noted that the sensitivity was based on 1000 unique combinations of the parameters. The sensitivity of most of the parameters varied with the different objective functions. The deep percolation fraction (RCHRGDP) is clearly identified as having by far the largest effect on the objective functions. It is the most sensitive parameter for the RMSE of nondriven flows and the second most sensitive for the RMSE of driven flows. Even in this simple case, however, it is not entirely clear how to interpret the sensitivity measurement where there are two objectives. Furthermore, the objectives themselves interact as shown below with the use of Bayesian networks.

Table 1. Parameter sensitivity measured by one at a time (OAT) perturbation.

PARAMETER	Description	OAT Sensitivity *	
		RMSE driven	RMSE nondriven
HRUSLP_FLAT	Subbasin/HRU slope, flat areas (< 5%)	0.738222 (1)	0.029085 (8)
RCHRGDP	Deep percolation fraction	0.333860 (2)	0.783473 (1)
AWHC	Available water holding capacity	0.169391 (3)	0.050535 (4)
GWREVAP	Groundwater revap coefficient	0.072713 (4)	0.108941 (2)
ESCO_RYER	Soil evaporation compensation factor for perennial grass	0.030201 (5)	0.038923 (5)
HRUSLP_MID	Subbasin/HRU slope, medium slope areas (5-20%)	0.020693 (6)	0.031537 (7)
ESCO_FRST	Soil evaporation compensation factor for forested areas	0.017529 (7)	0.012107 (10)
HRUSLP_STEEP	Subbasin/HRU slope, steep areas (> 20%)	0.017255 (8)	0.031961 (6)
CN2_RYER_C	Curve number for perennial grass in hydrologic group C	0.008515 (9)	0.009164 (11)
REVAP_MN	Shallow aquifer H <sub>2</sub> O threshold depth for revap or deep percolat.	0.008363 (10)	0.054043 (3)
GWQMN	Shallow aquifer H <sub>2</sub> O threshold depth for return flow to occur	0.006390 (11)	0.014738 (9)
ALPHABF	Baseflow alpha factor	0.005003 (12)	0.008391 (12)
GWDELAY	Groundwater delay time	0.002788 (13)	0.006619 (13)
CN2_FRST_B	Curve number for forest in hydrologic group B	0.001850 (14)	0.001831 (15)
SURLAG	Surface runoff lag coefficient	0.001432 (15)	0.000249 (23)
CN2_RYER_D	Curve number for perennial grass in hydrologic group D	0.001110 (16)	0.003227 (14)
SMFMN	Melt factor for snow on December 21	0.000973 (17)	0.000446 (17)
SMTMP	Snowfall temperature	0.000634 (18)	0.000357 (20)
CN2_FRST_D	Curve number for forest in hydrologic group D	0.000538 (19)	0.000393 (18)
SLSUB_STEEP	Slope length, steep areas (> 20%)	0.000329 (20)	0.001715 (16)
CN2_FRST_C	Curve number for forest in hydrologic group C	0.000323 (21)	0.000280 (22)
SLSUB_MID	Slope length, medium slope areas (> 5-20%)	0.000257 (22)	0.000293 (21)
CHK2	Effective hydraulic conductivity in main channel alluvium	0.000149 (23)	0.000358 (19)
SMFMX	Melt factor for snow on June 21	0.000046 (24)	0.000010 (24)
SLSUB_FLAT	Slope length, flat areas (< 5%)	0.000030 (25)	0.000006 (25)

\*Numbers in parenthesis denote OAT sensitivity rank (lower means more sensitive).

## Bayesian networks and parameter interdependence

Although the Bayesian networks are commonly used to assess causality, in this study, the network was used to show how the genetic algorithm reached parameter values for a population on the Pareto optimal front. This approach contrast with previous automatic model calibrations

where a calibrated model is computed using one of the available techniques and the end result is known, but the way that the calibration algorithm used the parameters to achieve an optimum calibration is unknown. As shown in Figure 3, the network does well in the simulation of the Pareto front, implying that the network correctly estimated the relationships of the model parameters. The relation of the network variance about the calculated front to sources of error requires further research. We speculate that the variance is more meaningful than a simple artifact of the statistical assumptions of the Bayesian network.

Figure 5 (event driven subset), Figure 6 (equal weight subset), and Figure 7 (nondriven subset) show the networks learned by application of the Bayesian network. These networks should not be interpreted as a description of physical causality. In the center of Figure 5, the network has a parameter (node) for the RCHRGDP linked to ALPHABF, SMFMN, SLSUB\_STEEP, and CN2\_RYER\_D as its parents. There is no physical causality between these parameters in the SWAT model. What the link means is that a change in the parameters ALPHABF, SMFMN, SLSUB\_STEEP, and CN2\_RYER\_D requires a change in the parameter RCHRGDP. The assumption of global parameter independence and the requirement of an acyclic graph are evident in this interpretation. The node RCHRGDP is a Gaussian regression function of ALPHABF, SMFMN, SLSUB\_STEEP, and CN2\_RYER\_D. CN2\_RYER\_D has no parents and is described by the mean and variance of the data, while the other three parents are connected to other parameters. The regression equation is:  $RCHRGDP = -0.460 + 0.165ALPHABF + 0.004CN2\_RYER\_D - 0.006SLSUB\_STEEP + 0.014SMFMN$ ; with a variance of 0.00014. It is also possible that a parameter has an indirect and direct connection to the other nodes (e.g., CN2\_RYER\_D and RCHRGDP). In the same manner, each parameter used in the calibration is expressed as a regression equation with intercept, variance, and the parent parameters as coefficients. It should be noted that the nodes that are not directly connected to the objective functions have an indirect effect on the objective functions.

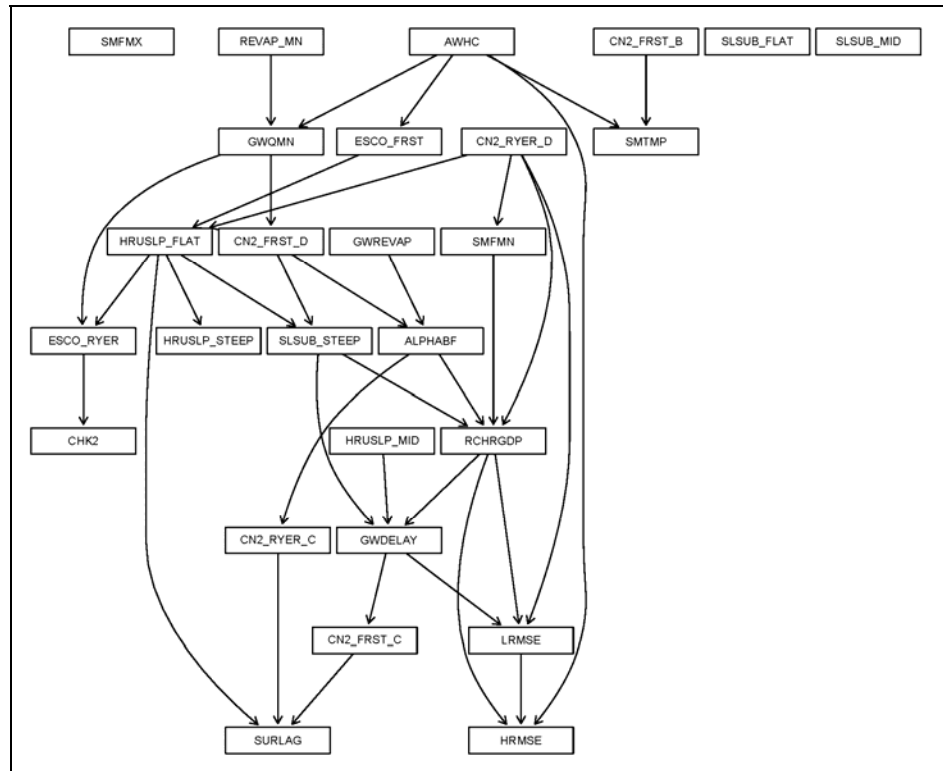


Figure 5. Bayesian network estimated for the event driven subset of calibrated SWAT models.

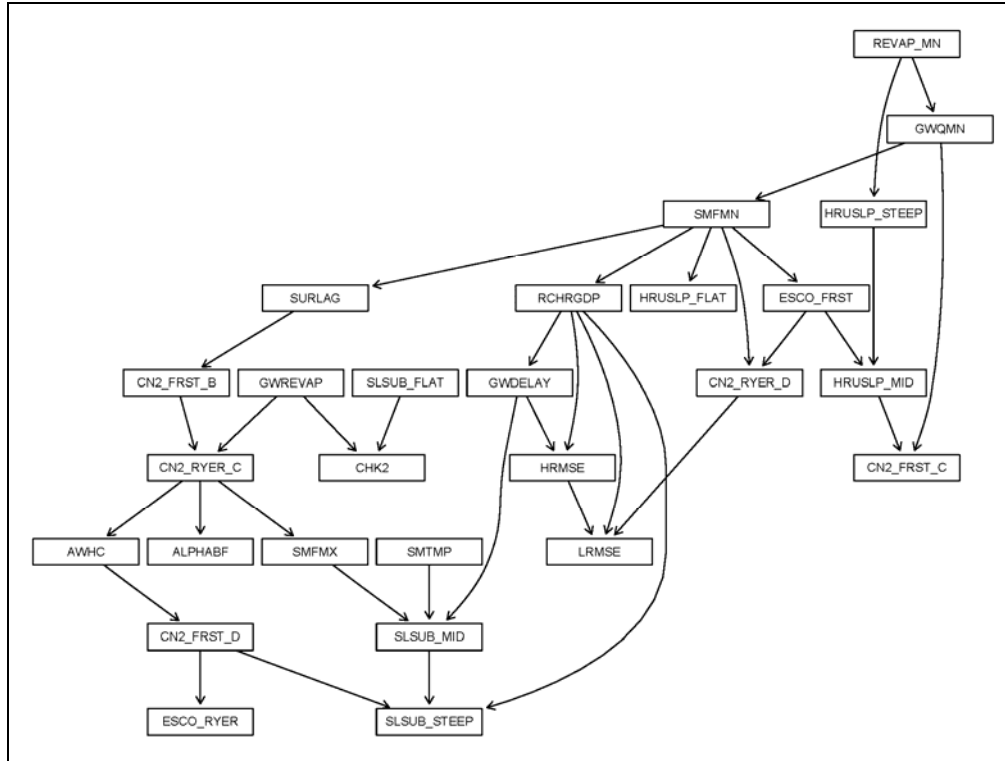


Figure 6. Bayesian network estimated for the equal weight subset of calibrated SWAT models.

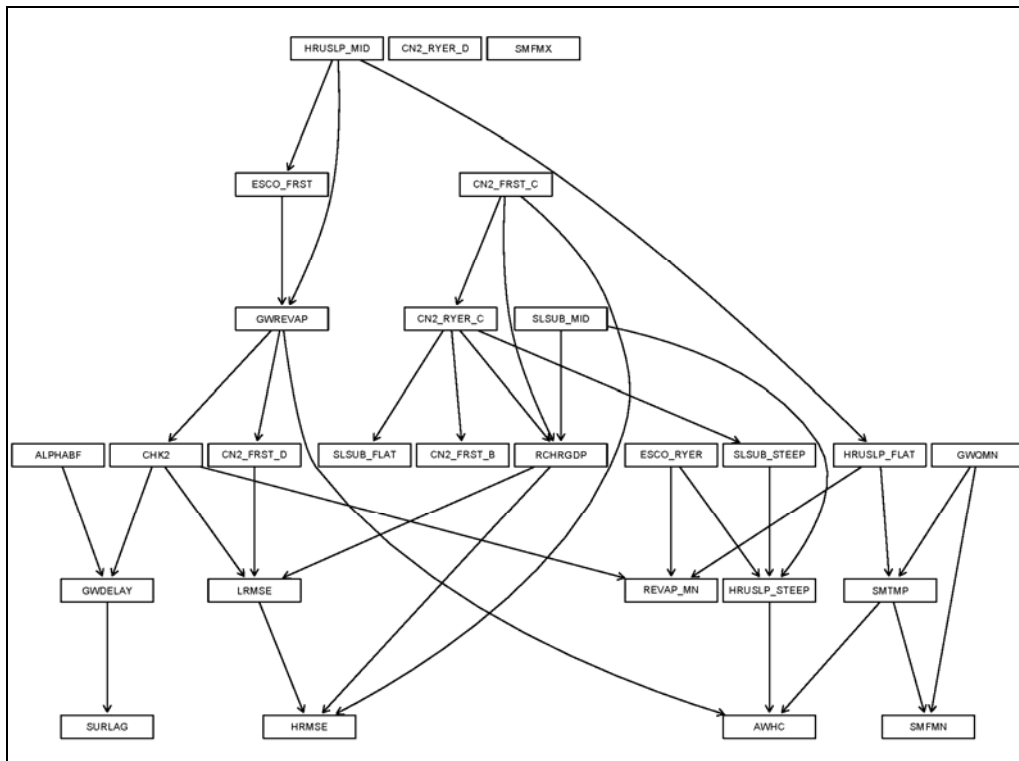


Figure 7. Bayesian network estimated for the nondriven subset of calibrated SWAT models.

From the networks, it is evident that the parameters and their interaction change with position along the Pareto front. It is interesting that RCHRGDP was the most OAT-sensitive parameter and had a direct link to the objective functions in all the networks. However, other OAT-sensitive parameters (AWHC, GWREVAP, HRU\_FLAT) were directly and indirectly connected to the objective functions in the event-driven network, and not connected to the objective outputs in the other networks. Furthermore, CN2\_RYER\_D was not OAT-sensitive but was directly connected to the two objective functions in the event-driven network; directly connected to the RMSE of nondriven flow in the equal weight network; and not connected at all in the nondriven network. It is noteworthy that the network changes with movement of the front as the autocalibration computation moves toward convergence. We hypothesize that there is useful information in these changes and will further explore this topic in future studies.

## **Conclusions**

The objective of the study was to analyze the effect of parameters used in multiobjective automatic calibration of a hydrologic model. Information about the interaction of parameters in the calibration was not available from the application of previously published methods for sensitivity measurement. Application of a Bayesian network method to the results of the calibration provided information on the interaction of parameters used by the calibration algorithm to find Pareto optimum calibrated models. Where there are multiple objectives, the parameters and their interaction in searching for the Pareto optimum change with position along the Pareto front. The information available from a Bayesian network requires redefining sensitivity to include a description of the interaction of parameters in the calibration search process.

## **Acknowledgements**

This study was funded by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) through the Conservation Effects Assessment Project (CEAP).

## **References**

- Arnold, J.G. and P.M. Allen. 1999. Automated methods for estimating baseflow and ground water recharge from streamflow records. *Journal of the American Water Resources Association* 35(2): 411-424.
- Arnold, J.G., P.M. Allen, R. Muttiah, and G. Bernhardt. 1995. Automated base flow separation and recession analysis techniques. *Ground Water* 33(6): 1010-1018.
- Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams. 1998. Large area hydrologic modeling and assessment part I: model development. *Journal of the American Water Resources Association* 34(1): 73-89.
- Beven, K. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249(1-4): 11-29.
- Borsuk, M.E., P. Reichert, A. Peter, E. Schager, and P. Burkhardt-Holm. 2006. Assessing the decline of brown trout (*Salmo trutta*) in Swiss rivers using a Bayesian probability network. *Ecological Modelling* 192(1-2): 224-244.

- Bøttcher, S.G. 2001. Learning Bayesian networks with mixed variables. In *Proc. of Eighth International Workshop on Artificial Intelligence and Statistics*, 149-156. T. Jaakkola and T. Richardson, eds. Key West, Florida: Morgan Kaufmann Publishers.
- Bøttcher, S.G. and C. Dethlefsen. 2003. Deal: A package for learning bayesian networks. *Journal of Statistical Software* 8(20): 1-40.
- Boyle, D.P., H.V. Gupta, S. Sorooshian, V. Koren, Z. Zhang, M. Smith. 2001. Toward improved streamflow forecasts: value of semidistributed modeling. *Water Resources Research* 37(11): 2749 - 2759.
- Confesor, R.B. Jr. and G.W. Whittaker. 2007. Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and pareto optimization. Accepted, *Journal of the American Water Resources Association*.
- Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2): 182-197.
- Di Luzio, M., Srinivasan, R., Arnold, J.G. and Neitsch, S.L., 2002. ArcView interface for SWAT2000, User's Guide.
- Eckhardt, K. and Arnold, J.G., 2001. Automatic calibration of a distributed catchment model. *Journal of Hydrology* 251(1-2): 103-109.
- Feyen, L., J.A. Vrugt, B.Ó. Nualláin, J. Knijff, and A. De Roo. 2006. Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology* 332(3-4): 276-289.
- Ho, C.M., R.A. Cropp, and R.D. Braddock. 2005. On the sensitivity analysis of two hydrologic models. In *Proc. MODSIM 2005 International Congress on Modelling and Simulation*, 2491-2497. A. Zerger and R.M. Argent, eds. Modelling and Simulation Society of Australia and New Zealand. Available at [http://www.mssanz.org.au/modsim05/papers/ho\\_cm.pdf](http://www.mssanz.org.au/modsim05/papers/ho_cm.pdf). Accessed April 4, 2007.
- Lauritzen, S.L., 2003. Some modern applications of graphical models. In *Highly structured stochastic systems*, 13-28. P.J. Green, S. Richardson, and N.L. Hjort, eds. Oxford: Oxford University Press.
- Loague, K.M. And Freeze, R.A. 1985. A comparison of rainfall runoff modeling techniques on small upland catchments. *Water Resources Research* 21(2): 229-248.
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models-part I: A discussion of principles. *Journal of Hydrology* 10(3): 282-290.
- Neitsch, S.L., J.G. Arnold, J.R. Kiniry, R. Srinivasan, J.R. Williams. 2002. Soil and water assessment tool user's manual: Version 2000. Agricultural Research Service, USDA, Temple, Texas. Available at <http://www.brc.tamus.edu/swat/downloads/doc/swatuserman.pdf>. Accessed April 10, 2007.
- Tarantola, S. and A. Saltelli. 2003. SAMO 2001: Methodological advances and innovative applications of sensitivity analysis. *Reliability Engineering & System Safety* 79(2): 121-122.
- Jasper A. Vrugt, H. V. Gupta, S. C. Dekker, S. Sorooshian, T. Wagener and W. Bouten. 2006. Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting model. *Journal of Hydrology* 325(1-4): 288-307.
- Van Liew, M. W., J. G. Arnold, and D. D. Bosch. 2005. Problems and potential of autocalibrating a hydrological model. *Transactions of the ASAE* 48(3): 1025-1040.